

Matrix Kummer-Pearson VII Relation and its Application in Affine Shape

Francisco J. Caro-Lopera

Department of Basic Sciences
Universidad de Medellín
Carrera 87 No.30-65, of. 5-103
Medellín, Colombia
E-mail: fjcaro@udem.edu.co

José A. Díaz-García *

Department of Statistics and Computation
Universidad Autónoma Agraria Antonio Narro
25350 Buenavista, Saltillo, Coahuila, Mexico
E-mail: jadiaz@uaaan.mx

Abstract

A case of the matrix Kummer relation of Herz (1955) based on the Pearson VII type matrix model is derived in this paper. As a consequence, the polynomial Pearson VII configuration density is obtained and this set the corresponding exact inference as a solvable aspect in shape theory.

1 Introduction

Start assuming that a given sample of n “figures”, comprised in N landmarks (“anatomical” points) in K dimensions, and summarized in an $N \times K$ matrix \mathbf{X} , belongs to certain matrix variate distribution with unknown scale and location parameters. The statistical theory of shape pursues the distribution of the transforming \mathbf{X} after filtering out some non important geometrical aspect of the original figures, such as the scale, the position, the rotation, the uniform shear, and so on; then the so called shape of the population object should be inferred and some comparisons among population shapes could be performed, among many others statistical comparisons in the shape space, instead of comparisons in the very noised non transformed Euclidean space. As usual, the classical theory of shape, with different geometrical filters, as the Euclidean or affine, was based on Gaussian samples, see for example Goodall and Mardia (1993) and the references therein. Then, the non normal applications demanded general assumptions for the samples, and the so termed generalized shape theory was set under elliptical models. Under the Euclidean filter we can mention the work of Díaz-García and Caro-Lopera (2010), Díaz-García and Caro-Lopera (2012); they find the density of all geometrical information about the elliptical random \mathbf{X} which remains after removing the scale, the position and the rotation. Finally, if an affine filter is applied,

*Corresponding author

Key words. Matrix Kummer relation, Pearson VII configuration density, zonal polynomials.
2000 Mathematical Subject Classification. Primary 62E15; 60E05; secondary 62H99

in order to remove from \mathbf{X} all geometrical information of scale, position and uniform shear, Caro-Lopera *et al* (2010) obtained the so term configuration density of \mathbf{X} . All the above densities are expanded in terms of the well known zonal polynomials, in a series of papers by A.T. James in 60's, see for example Muirhead (1982).

The transition of the Gaussian shape theory to the elliptical shape theory demanded some advances in integration involving zonal polynomials (see for example Caro-Lopera *et al* (2010)), but important problems remain, the computability of the shape densities.

In this paper we focus in alternatives for such problems. It is easy to see in Goodall and Mardia (1993), Díaz-García and Caro-Lopera (2010) and Díaz-García and Caro-Lopera (2012) that the structure of shape densities under Euclidean transformations involves series of zonal polynomials which heritages the difficulties for computations of the classical hypergeometric series studied by Koev and Edelman (2006). However, a class of the generalized confluent type series of the configuration densities of Caro-Lopera *et al* (2010) can be handled in order to transformed the series into polynomials, and then the addressed open problems for computations of the series can be avoided.

The configuration density under Kotz type samples (including Gaussian) has this property, and the inference can be performed with polynomials instead of infinite series, see Caro-Lopera *et al* (2009). The source for this property resides in a generalization of the Kummer relation of Herz (1955).

This motivates the present work, claiming that there is a similar Kummer relation based on a Pearson VII distribution, which under certain restriction of the parameters in the associated Pearson VII configuration density, it can be turned in a polynomial density. Then the inference can be performed easily by working with the exact likelihood which is written in terms of very low zonal polynomials. In fact the exact densities can be write down by using formulae for those polynomials; for example, in the planar shape theory (the most classical applications resides in the study of figures in \mathbb{R}^2), we can use directly the formulae given by Caro-Lopera *et al* (2007) instead of the numerical approaches by Koev and Edelman (2006) in order to perform some analytical properties of the exact density.

This discussion is placed in this paper as follows: section 2 defines a Pearson VII type series and finds an integral representation which lead to a matrix Kummer type relation which we call Kummer-Pearson VII relation; then by applying some general properties studied by Herz (1955), the equality is extended for the required domains in the shape theory context. Finally, Kummer-Pearson VII relation gives the finiteness of the Pearson VII configuration density in section 3.

2 Matrix Kummer-Pearson VII relation

Recall that the matrix Kummer relation (due to Herz (1955)) states that

$${}_1F_1(a; c; \mathbf{X}) = \text{etr}(\mathbf{X}) {}_1F_1(c - a; c; -\mathbf{X}) \quad \text{and}; \quad (1)$$

see also Muirhead (1982).

Now, let $\mathbf{X} > 0$ be an $m \times m$ positive definite matrix, then define

$${}_1P_1(f(t, \mathbf{X}) : a; c; \mathbf{X}) = \sum_{t=0}^{\infty} \frac{f(t, \mathbf{X})}{t!} \sum_{\tau} \frac{(a)_{\tau}}{(c)_{\tau}} C_{\tau}(\mathbf{X}), \quad (2)$$

where the function $f(t, \mathbf{X})$ is independent of τ , $\tau = (t_1, \dots, t_m)$, $t_1 \geq t_2 \geq \dots \geq t_m > 0$, is a partition of t ,

$$(\beta)_{\tau} = \prod_{i=1}^m \left(\beta - \frac{1}{2}(i-1) \right)_{t_i}$$

and

$$(b)_t = b(b+1) \cdots (b+t-1), \quad (b)_0 = 1.$$

Then, using this notation we see that the Kummer relation (1) is a particular case of a general type of expressions with the following form

$${}_1P_1(f(t, \mathbf{X}) : a; c; \mathbf{X}) = v(\mathbf{X}) {}_1P_1(g(t, \mathbf{X}) : c-a; c; h(\mathbf{X})), \quad (3)$$

where the functions v, g and h are uniquely determined by the particular function f and according to the domain of the parameters a and c and the matrix \mathbf{X} .

First, we consider an integral representation of the left hand side of (3) under the model $f(t, \mathbf{X}) = (b)_t$.

Theorem 2.1. *Let $\mathbf{X} < \mathbf{I}$, $\text{Re}(a) > \frac{1}{2}(m-1)$, $\text{Re}(c) > \frac{1}{2}(m-1)$ and $\text{Re}(c-a) > \frac{1}{2}(m-1)$. Then for suitable reals b and d , we have that*

$$\begin{aligned} {}_1P_1((b)_t d^{-b-t} : a; c; \mathbf{X}) &= \frac{\Gamma_m(c)}{\Gamma_m(a)\Gamma_m(c-a)} \\ &\times \int_{0 < \mathbf{Y} < \mathbf{I}_m} (d - \text{tr}(\mathbf{XY}))^{-b} |\mathbf{Y}|^{a-(m+1)/2} |\mathbf{I} - \mathbf{Y}|^{c-a-(m+1)/2} (d\mathbf{Y}). \end{aligned} \quad (4)$$

Proof. First, we use a zonal polynomial expansion

$$\begin{aligned} (d - \text{tr}(\mathbf{XY}))^{-b} &= \sum_{t=0}^{\infty} \frac{(b)_t d^{-b-t}}{t!} [\text{tr}(\mathbf{XY})]^t \\ &= \sum_{t=0}^{\infty} \frac{(b)_t d^{-b-t}}{t!} \sum_{\tau} C_{\tau}(\mathbf{XY}). \end{aligned}$$

Then integrating term by term using Muirhead (1982, theorem 7.2.10), we have that

$$\begin{aligned} &\int_{0 < \mathbf{Y} < \mathbf{I}_m} (d - \text{tr}(\mathbf{XY}))^{-b} |\mathbf{Y}|^{a-(m+1)/2} |\mathbf{I} - \mathbf{Y}|^{c-a-(m+1)/2} (d\mathbf{Y}) \\ &= \sum_{t=0}^{\infty} \frac{(b)_t d^{-b-t}}{t!} \sum_{\tau} \int_{0 < \mathbf{Y} < \mathbf{I}_m} |\mathbf{Y}|^{a-(m+1)/2} |\mathbf{I} - \mathbf{Y}|^{c-a-(m+1)/2} C_{\tau}(\mathbf{XY}) (d\mathbf{Y}) \\ &= \sum_{t=0}^{\infty} \frac{(b)_t d^{-b-t}}{t!} \sum_{\tau} \frac{(a)_{\tau}}{(c)_{\tau}} \frac{\Gamma_m(a)\Gamma_m(c-a)}{\Gamma_m(c)} C_{\tau}(\mathbf{X}) \\ &= \frac{\Gamma_m(a)\Gamma_m(c-a)}{\Gamma_m(c)} \sum_{t=0}^{\infty} \frac{(b)_t d^{-b-t}}{t!} \sum_{\tau} \frac{(a)_{\tau}}{(c)_{\tau}} C_{\tau}(\mathbf{X}) \\ &= \frac{\Gamma_m(a)\Gamma_m(c-a)}{\Gamma_m(c)} {}_1P_1((b)_t d^{-b-t} : a; c; \mathbf{X}), \end{aligned}$$

and the required result follows. \square

Now, we derive the version of (1) but based on a Pearson VII type model, we call this expression, Kummer-Pearson VII relation.

Theorem 2.2. *Let $\mathbf{X} > 0$, $\text{Re}(a) > \frac{1}{2}(m-1)$, $\text{Re}(c) > \frac{1}{2}(m-1)$ and $\text{Re}(c-a) > \frac{1}{2}(m-1)$. Then for suitable reals b and d , the Kummer-Pearson VII relation is given by*

$${}_1P_1((b)_t d^{-b-t} : a; c; \mathbf{X}) = (d - \text{tr} \mathbf{X})^{-b} {}_1P_1\left((b)_t (d - \text{tr} \mathbf{X})^{-t} : c-a; c; -\mathbf{X}\right). \quad (5)$$

Proof. Consider $\mathbf{W} = \mathbf{I} - \mathbf{Y}$ in (4), then we obtain

$$\begin{aligned}
{}_1P_1((b)_t d^{-b-t} : a; c; \mathbf{X}) &= \frac{\Gamma_m(c)}{\Gamma_m(a)\Gamma_m(c-a)} \\
&\times \int_{0 < \mathbf{W} < \mathbf{I}_m} (d - \text{tr}[\mathbf{X}(\mathbf{I} - \mathbf{W})])^{-b} |\mathbf{W}|^{c-a-(m+1)/2} |\mathbf{I} - \mathbf{W}|^{a-(m+1)/2} (d\mathbf{W}) \\
&= \frac{\Gamma_m(c)}{\Gamma_m(a)\Gamma_m(c-a)} \\
&\times \int_{0 < \mathbf{W} < \mathbf{I}_m} (d - \text{tr} \mathbf{X} - \text{tr}(-\mathbf{X}\mathbf{W}))^{-b} |\mathbf{W}|^{c-a-(m+1)/2} |\mathbf{I} - \mathbf{W}|^{a-(m+1)/2} (d\mathbf{W}) \\
&= \frac{\Gamma_m(c)}{\Gamma_m(a)\Gamma_m(c-a)} \\
&\times \frac{\Gamma_m(a)\Gamma_m(c-a)}{\Gamma_m(c)} {}_1P_1((b)_t (d - \text{tr} \mathbf{X})^{-b-t} : c-a; c; -\mathbf{X}),
\end{aligned}$$

which is the required result. \square

The reader can compare theorem 2.2 (and its proof) with the Kummer relation (and its proof given by Herz (1955)). So the analysis of Herz (1955) for extending the above relations for other values of the parameters, holds in the Kummer-Pearson VII relation too.

Explicitly, we proved that the integral representation (4) of ${}_1P_1((b)_t : a; c; \mathbf{X})$ holds for $\text{Re}(\mathbf{X}) < \mathbf{I}$ (by analytic continuation), $\text{Re}(a) > \frac{1}{2}(m-1)$, $\text{Re}(c) > \frac{1}{2}(m-1)$ and $\text{Re}(c-a) > \frac{1}{2}(m-1)$. Then by a suitable modification of the arguments in Herz (1955), we can extend the domain of (5) as follows

Theorem 2.3. *$\text{Re}(\mathbf{X}) > 0$, $\text{Re}(a) > \frac{1}{2}(m-1)$ and $\text{Re}(c) > \frac{1}{2}(m-1)$. Then for suitable complex numbers b and d , the Kummer-Pearson VII relation is given by*

$${}_1P_1((b)_t d^{-b-t} : a; c; \mathbf{X}) = (d - \text{tr} \mathbf{X})^{-b} {}_1P_1((b)_t (d - \text{tr} \mathbf{X})^{-t} : c-a; c; -\mathbf{X}). \quad (6)$$

The above relation is important in shape theory applications, in the so called *polynomial Pearson VII configuration density*.

3 Polynomial Pearson VII Configuration Density

Our motivation for studying finite shape densities, comes from the computations of hypergeometric series type involved in these distributions. It is known, that the zonal polynomials are computable very fast by Koev and Edelman (2006), but the problem now resides in the convergence and the truncation of the series of zonal polynomials. In fact, in the same reference of Koev and Edelman (2006) we read:

“Several problems remain open, among them automatic detection of convergence and it is unclear how to tell when convergence sets in. Another open problem is to determine the best way to truncate the series. ”

Thus the implicit numerical difficulties for truncation of any configuration density motivate two areas of investigation: first, continue the numerical approach started by (Koev and Edelman (2006)) with the confluent hypergeometric functions and extend it to the case of some configuration series type Kotz, Pearson VII, Bessel, Logistic, for example; or second, propose a theoretical approach for solving the problem analytically (see Caro-Lopera *et al* (2009)).

We study now the second question corresponding to the polynomial Pearson VII configuration density.

Recall that a $p \times n$ random matrix \mathbf{X} is said to have a matrix variate symmetric Pearson type VII distribution with parameters $s, R \in \mathfrak{R}, M : p \times n, \Sigma : p \times p, \Phi : n \times n$ with $R > 0, s > np/2, \Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{\Gamma(s)}{(\pi R)^{np/2} \Gamma(s - \frac{np}{2}) |\Sigma|^{n/2} |\Phi|^{p/2}} \left(1 + \frac{\text{tr}(\mathbf{X} - \mathbf{M})' \Sigma^{-1} (\mathbf{X} - \mathbf{M}) \Phi^{-1}}{R} \right)^{-s}.$$

When $s = (np + R)/2$, \mathbf{X} is said to have a matrix variate t -distribution with R degrees of freedom. And in this case, if $R = 1$, then \mathbf{X} is said to have a matrix variate Cauchy distribution, see Caro-Lopera *et al* (2010).

Then, by Caro-Lopera *et al* (2010) we have that (see Goodall and Mardia (1993) for the gaussian case),

Lemma 3.1. *Let be*

$$A = \frac{\Gamma_K(\frac{N-1}{2})}{\pi^{Kq/2} |\Sigma|^{\frac{K}{2}} |U' \Sigma^{-1} U|^{\frac{N-1}{2}} \Gamma_K(\frac{K}{2})}, \quad a = \frac{N-1}{2} \quad (7)$$

$$\mathbf{X} = \frac{1}{R} U' \Sigma^{-1} \mu \mu' \Sigma^{-1} U (U' \Sigma^{-1} U)^{-1}, \quad c = \frac{K}{2}, \quad (8)$$

$$b = s - \frac{K(N-1)}{2}, \quad d = 1 + \frac{\text{tr}(\mu' \Sigma^{-1} \mu)}{R}. \quad (9)$$

If $\mathbf{Y} \sim E_{N-1 \times K}(\mu_{N-1 \times K}, \Sigma_{N-1 \times N-1} \otimes \mathbf{I}_K, h)$, for $\Sigma > 0$, then the non-isotropic noncentral Pearson type VII configuration density is given by

$$A_1 P_1((b)_t d^{-b-t} : a; c; \mathbf{X}), \quad (10)$$

where ${}_1P_1(\cdot)$ has been defined in (2).

Unfortunately, the above configuration density with general form $A_1 P_1(f(t) : a; c; \mathbf{X})$ is an infinite series, given that $a = \frac{N-1}{2}$ and $c = \frac{K}{2}$ are positive (recall that N is the number of landmarks, K is de dimension and $N - K - 1 \geq 1$). So a truncation is needed for performing inference when the modified algorithms of Koev and Edelman (2006) are used.

But the above series can be turned into a polynomial if we use the following basic principle of Caro-Lopera *et al* (2009).

Lemma 3.2. *Let $N - K - 1 \geq 1$ as usual, and consider the definition of ${}_1P_1(\cdot)$ in (2). The infinite configuration density has the general form*

$$CD_1 = w(K, N, \mathbf{X}) {}_1P_1\left(f(t, \mathbf{X}) : \frac{N-1}{2}; \frac{K}{2}; \mathbf{X}\right),$$

for suitable functions: $w(\cdot)$, independent of t and τ but dependent of K, N and \mathbf{X} ; and $f(\cdot)$, independent of τ , but dependent of t and possibly of \mathbf{X} (it depends on the generator elliptical function, compare with the particular Pearson VII case of lemma 3.1). Then, according to (3), if the dimension K is even (odd) and the number of landmarks N is odd (even), respectively, then the equivalent configuration density

$$CD_2 = w(K, N, \mathbf{X}) v(\mathbf{X}) {}_1P_1\left(g(t, \mathbf{X}) : -\left(\frac{N-1}{2} - \frac{K}{2}\right); \frac{K}{2}; h(\mathbf{X})\right)$$

is a polynomial of degree $K(\frac{N-1}{2} - \frac{K}{2})$ in the latent roots of the matrix \mathbf{X} (otherwise the series is infinite); where v, g and h are functions understood in the context of (3) and depends on the elliptical generator function.

Given an elliptical configuration density CD_1 indexed by the function $f(\cdot)$ and based in the fact that $a = \frac{N-1}{2} > 0, c = \frac{K}{2} > 0$, the crucial point here consists of finding an integral representation valid for $c - a = -\frac{N-K-1}{2} < 0$, which will lead to an equivalent elliptical configuration density CD_2 indexed by some function $g(\cdot)$. Then the finiteness of CD_2 follows from K even (odd) and N odd (even), respectively.

In particular, for the Pearson VII generator function, the referred polynomial density is provided by applying the new Kummer-Pearson VII relation of theorem 2.3, via lemma 3.2, in lemma 3.1.

Theorem 3.1. *Let be A, a, b, c, d and \mathbf{X} defined by (7)-(9).*

If $\mathbf{Y} \sim E_{N-1 \times K}(\mu_{N-1 \times K}, \Sigma_{N-1 \times N-1} \otimes \mathbf{I}_K, h)$, $\Sigma > 0$, K is even (odd) and N is odd (even), respectively, then the polynomial non-isotropic noncentral Pearson type VII configuration density is given by

$$(d - \text{tr } \mathbf{X})^{-b} {}_1P_1 \left((b)_t (d - \text{tr } \mathbf{X})^{-t} : c - a; c; -\mathbf{X} \right) \quad (11)$$

and it is a polynomial of degree $K \left(\frac{N-1}{2} - \frac{K}{2} \right)$ in the latent roots of \mathbf{X} .

Proof. The proof is trivial, just start with the infinite configuration density (10):

$$A {}_1P_1 \left((b)_t d^{-b-t} : a; c; \mathbf{X} \right),$$

where A, a, b, c, d and \mathbf{X} are given by (7)-(9). Then apply (6) and the result follows. Note that the finiteness follows from Lemma (3.2) by noting that $c - a = -\frac{N-K-1}{2}$ is a negative integer, when K is even (odd) and N is odd (even). \square

The principle of lemma 3.2 is based on a known property of the hypergeometric series easily extended to series of the type ${}_1P_1(f(t, \mathbf{X}) : a, c; \mathbf{X})$, see (2), which states that if a is a negative integer or a negative half integer, the series vanishes in a polynomial. Then if we have an application following a confluent distribution type, it is a polynomial, always that the parameter a accepts the addressed special domain, otherwise the distribution is a series of zonal polynomials and the open problems for its computability appears. This last case occurs for example in the general configuration density, which is an infinite confluent series type because $a = \frac{N-1}{2} > 0$, then it is not a trivial to turn that series into polynomials, because the application does not allow a negative parameter. Then the lemma 3.2 gives a special subclass by selecting the number of landmarks (even, odd) given the dimension (odd, even), otherwise the configuration density remains a series, then the associated Kummer relation must be obtained in order to transform the numerator parameter a into $c - a$ which is a negative half integer as required. So, in the context of shape theory under affine transformations, the study of Kummer relation type plays an important role. It is easy to check that the classical Kummer relation first derive by Herz (1955), and set in the context of zonal polynomials by Constantine (1963) is related with a Gaussian kernel, but if we want to obtain some non Gaussian polynomial densities (under the explained restrictions of N and K), then we need to derive new Kummer relations, it was the case of certain class of Kotz configuration densities (which includes the classical Gaussian), it need the derivation of the associated Kummer Kotz relation, see Caro-Lopera *et al* (2009). Thus, in the case of the Pearson polynomial configuration density the corresponding Kummer relation was the key point for transforming series.

The above discussion opens related problems in some special topics of matrix variate analysis involving confluent matrix with special domains for the parameters, or general hypergeometric series type with more or equal Pochhammer symbols in the numerator,

several examples of this situations, which demands new developments for Euler relations and similar ones, can be inferred from some distributions proposed in Muirhead (1982, chapters 8–11).

Acknowledgment

The first author was supported by Universidad de Medellín-Colombia grant No. 158.

References

- F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2007). A formula for Jack polynomials of the second order. *Appl. Math.* 34, 113–119.
- F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías.(2009). Inference in statistical shape theory: elliptical configuration densities. *Journal of Statistical Research*, Vol. 43(1), 1–19.
- F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2010). Noncentral Elliptical Configuration Density. *Journal of Multivariate Analysis*. 101(1), 32–43.
- A. G. Constantine. (1963). Noncentral distribution problems in multivariate analysis. *The Annals of Mathematical Statistics*, 34,1270–1285.
- J. A. Díaz-García and F. J. Caro-Lopera. (2010). Generalised shape theory via SVD decomposition I. *Metrika*.DOI:10.1007/s00184-010-0341-5. Online First.
- J. A. Díaz-García and F. J. Caro-Lopera. (2012). Statistical theory of shape under elliptical models and singular value decompositions. *Journal of Multivariate Analysis*.103,77-92.
- C. R. Goodall, and K.V. Mardia. (1993). Multivariate Aspects of Shape Theory. *Ann. Statist.*, 21, 848–866.
- C. S. Herz. (1955). Bessel Functions of Matrix Argument. *Ann. Math.*, 61, 474–523.
- P. Koev and A. Edelman. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. *Math. Comp.* 75, 833-846.
- R. J. Muirhead. (1982). *Aspects of multivariate statistical theory*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc.,